Running head: Estimating length of stay

Short title: Estimating the effect of MRSA infection on length of stay using a longitudinal model

Abbreviations: acute physiology and chronic health evaluation (APACHE), credible interval (CI), deviance information criterion (DIC), intensive care unit (ICU), Markov chain Monte Carlo (MCMC), Methicillin-resistant *Staphylococcus aureus* (MRSA), Methicillin-sensitive *Staphylococcus aureus* (MSSA), relative risk ratio (RRR), systemic inflammatory response syndrome (SIRS), therapeutic intervention scoring system (TISS)

Word count: 200 (abstract), 3975 (text)

Abstract

Healthcare-associated methicillin-resistant *Staphylococcus aureus* (MRSA) infection may cause increased hospital stay or, sometimes, death. Quantifying this effect is complicated because it is a time-dependent exposure: infection may prolong hospital stay, while longer stays increase the risk of infection. We overcome these problems by using a multinomial longitudinal model for estimating the daily probability of death and discharge. We then extend the basic model to estimate how the effect of MRSA infection varies over time, and to quantify the number of excess ICU days due to infection. We find that infection decreases the relative risk of discharge (relative risk ratio = 0.68, 95% credible interval: 0.54, 0.82), but is only indirectly associated with increased mortality. An infection on the first day of admission resulted in a mean extra stay of 0.3 days (95% CI: 0.1, 0.5) for a patient with an APACHE II score of 10, and 1.2 days (95% CI: 0.5, 2.0) for a patient with an APACHE II score of 30. The decrease in the relative risk of discharge remained fairly constant with day of MRSA infection, but was slightly stronger closer to the start of infection. These results confirm the importance of MRSA infection in increasing ICU stay, but suggest that previous work may have systematically overestimated the effect size.

Medical Subject Headings (MeSH) (keywords): Competing risks, Intensive Care, Longitudinal analysis, Nosocomial Infections, Time-dependent

Healthcare associated infections affect 5–10% of acute-care patients in developed countries, and considerably more in developing nations (1). These infections are direct causes of patient morbidity and mortality, and are also believed to lead to increased hospital stays. Many infections are preventable by the use of interventions (2). Infections places an important—but poorly quantified—burden on health services. Quantifying excess hospital stay is essential for assessing how many bed days might be gained from prevention and subsequent health economic analyses that inform the allocation of resources to infection control programmes (3). The recent decision by the Centers for Medicare and Medicaid Services to stop re-imbursements to US hospitals for selected healthcare associated infections increases the need for a valid interpretation of the costs and benefits of infection control interventions (4).

Estimating additional length of hospital stay due to nosocomial infections creates a number of statistical challenges (5). The central difficulty arises from the fact that infections may increase the length of stay, and increased length of stay simultaneously increases the chance of infection (6). However, most standard regression analyses assume a one-way direction of causation from exposure (infection) to response (length of stay). Standard survival analysis of hospital stay data is also inappropriate because censoring of hospital stays due to death does not occur at random. Instead, the most severely ill will have the highest chance of dying and the lowest chance of being discharged on a given day. Such informative censoring violates the

assumptions of standard survival analyses, and can lead to very large biases if unaccounted for. A further problem is that factors that may predispose a patient to infection (such as use of invasive interventions) may also be independently associated with increased length of stay.

In this paper we aim to address all these challenges, and to use individual patient-level data to quantify the additional length of stay in an intensive care unit (ICU) due to infections with one of the most virulent and widely distributed nosocomial pathogens, methicillin-resistant *Staphylococcus aureus* (MRSA).

We make use of a multinomial regression model with a longitudinal data format. This format has the advantage of allowing the use of time-dependent exposures (i.e. exposures such as assisted ventilation that may change during a patient's stay). Time-dependent exposures can be more informative than time-independent exposures because they offer the chance to examine the order of exposure and response.

A further advantage of the longitudinal format is that the effect of time-dependent exposures can be allowed to change over time by using time-dependent covariates. If modelled with only time-*independent* covariates, a hospital acquired infection would be assumed to decrease the risk of discharge by the same amount regardless of the length of time the patient had already spent in the ward. In contrast, time-dependent covariates allow changes in the exposure effects over time. Thus, for example, the effect of a nosocomial infection on the chances of being discharged might be strongest on the actual

day of infection and then wane over time.

The use of time-dependent exposures has been extensively studied in the field of longitudinal analyses (7, chapter 12) and in this paper we apply some of this knowledge to length of stay data.

MATERIALS AND METHODS

Clinical setting

Data were collected from all patients admitted to two adjacent 15 bed intensive care units (ICUs) in a 1100 bed dual-site teaching hospital in the United Kingdom between January 1st 2002 and April 20th 2006. During the study period 60% of the patients were medical, 21% surgical and 19% cardiothoracic. Data comprised age, sex, date of admission and discharge to ICU, speciality, day 1 Acute Physiology and Chronic Heath Evaluation (APACHE)-II score, daily therapeutic intervention scoring system (TISS) score (8) which included measurements required to diagnose a systemic inflammatory response syndrome, dates of starting or stopping ventilation or hemofiltration, date of collection and culture of MRSA from all microbiological samples, and date of starting treatment with vancomycin or linezolid. Clinical samples were only taken when local or systemic infection was suspected. Further details on infection control and laboratory practice have been previously reported (9, 10).

An MRSA infection was considered to be present if (and only if) three

conditions were satisfied: 1) MRSA was isolated from a sterile or non-sterile clinical sample including a removed vascular catheter tip; 2) there was treatment with vancomycin or linezolid which were the only antibiotics used for initial treatment of suspected or proven MRSA infection, started between 1 day before and 3 days after the positive culture; and 3) there was a systemic inflammatory response syndrome (SIRS) response, requiring 2 of the following criteria present between 2 days before and 3 days after the positive MRSA culture: temperature < 36 °C or > 38 °C; heart rate > 90 bpm; respiratory rate > 20 breaths/min or $PaCO_2 < 32$ mmHg; white blood cell count $> 12{,}000$ or $< 4{,}000$ cells/mm^3 (11). Patients from whom MRSA was isolated from any site but who did not fulfil these additional criteria were considered to be colonised with MRSA.

Data quality control

Data quality control mechanisms included automated range, logic and date checks. The integrity of the data extraction process was validated for completeness and accuracy by manually comparing 5% of the electronic database with the original source data.

Exposures

We were motivated to fit a longitudinal model because the data have a number of important time-dependent exposures. The most important being the

presence of an MRSA infection, because the primary research question was the impact of such infections on length of stay. We carried forward the effect of MRSA over time, so once a patient had an MRSA infection their status was "yes" for all subsequent days until discharge. This is because we expect the effect of an infection to persist; how long the effect persists was one of the questions we addressed in the time-dependent covariate model. The list of exposures is in Table 1.

Ethics approval

The hospital ethics committee waived the need for patient consent and agreed to the use of this anonymised patient database for this study.

Statistical methods

Daily ICU data were used for all analyses. The table below shows a subset of the data for two subjects. Each of the n rows corresponds to one patient day on the ICU.

ICUID	Date	TISS	Sex	Outcome
3154	07 Feb 02	54	F	Stayed
3154	08 Feb 02	37	F	Stayed
3154	09 Feb 02	40	F	Stayed
3154	10 Feb 02	27	F	Discharged
3163	09 Feb 02	51	M	Stayed
3163	10 Feb 02	39	M	Stayed
3163	11 Feb 02	49	M	Stayed
3163	12 Feb 02	60	M	Died

Patient 3154 was admitted to the ICU on February 7th and discharged on the 10th, giving a length of stay of four days and a final outcome of "Discharged". Patient 3163 was admitted on February 9th and died on the 12th, giving a length of stay of four days and a final outcome of "Died". The TISS score (a measure of the level of patient care required) is a time-dependent exposure which changed from day-to-day (8). The nominal response variable "Outcome" describes each patient's day-to-day status (stayed, discharged or died). We assumed this nominal response had a multinomial distribution, and so examined the probability of "Stayed", "Discharged" and "Died" for the ith patient day ($i = 1, \ldots, n$), denoted as π_{i1}, π_{i2} and π_{i3} respectively (12, Chapter 8). We were interested in the association between these three probabilities and the exposures, hence we used a nominal logistic regression

(or multinomial) model defined as

$$\begin{aligned}
\hat{\pi}_{i1} &= \frac{1}{1+\exp(r_{i2})+\exp(r_{i3})}, \qquad i=1,\ldots,n,\\
\hat{\pi}_{ij} &= \frac{\exp(r_{ij})}{1+\exp(r_{i2})+\exp(r_{i3})}, \qquad i=1,\ldots,n,\ j=2,3,\\
r_{ij} &= \mathbf{x}_i^T\mathbf{b}_j, \qquad i=1,\ldots,n,\ j=2,3, \qquad (1)
\end{aligned}$$

where $\mathbf{x}_i$ is a set of exposures for row i of the data and $\mathbf{b}_j$ is the vector of parameters for outcome j $(b_{j1},\ldots,b_{jp})$, where $j=2$ represents ICU discharge and $j=3$ represents death. The above formulation satisfies the multinomial assumption that $\pi_{i1}+\pi_{i2}+\pi_{i3}=1$.

The exponential of b_{jk} gives the relative risk ratio, for a one unit increase in the values of covariate k, of being in category $j=2,3$ relative to category $j=1$ (staying), given that the other covariates are held constant. For example, a value of $\exp(b_{2,5})=2$ would mean that the relative risk of being discharged would be twice as likely than staying when covariate 5 is increased by one. A relative risk ratio is similar to an odds ratio, but is necessarily more complicated because of the multiple response categories.

This nominal logistic regression model can be thought of as a discrete time longitudinal survival model (13, Section 10.2.3) and is closely related to a competing risks model (14). To realise this, consider that the multinomial model estimates the probability of death or discharge on each day. Similarly, if we used a competing risks model we would estimate the probability of death or discharge in a short period of time. Previous work has shown the

similarity between a logistic longitudinal survival model (i.e., $j = 1, 2$) and Cox regression (15). We have extended this similarity by changing the logistic model to a nominal logistic model, and the Cox regression to a competing risks model.

One advantage of using a nominal logistic model with a longitudinal structure is the ability to incorporate random effects. These are useful for modelling heterogeneity and allow the model to account for some of the large unexplained variation in length of stay. We considered models with a random intercept for each patient admission, allowing the probability of discharge and death to vary between admissions. The regression equation (1) becomes

$$r_{ij} = \mathbf{x}_i^T \mathbf{b}_j + z_{sj} I(s_i = s), \qquad i = 1, \ldots, n,\ j = 2, 3,$$

where z_{sj} is the random intercept for admission s for discharge ($j = 2$) and death ($j = 3$). As before, we do not need to specify an intercept for the reference category ($j = 1$). The $I()$ is an indicator function which matches the admission number on row i of the data to admission s. We used a multivariate Normal distribution to create each admission's death and discharge intercept

$$\mathbf{z}_s \sim \mathbf{N}(\mathbf{0}, \Omega), \qquad s = 1, \ldots, m,$$

where Ω is a 2×2 variance–covariance matrix and m is the total number of subjects.

We also considered models with time-dependent covariates, allowing the effect of MRSA infection on subsequent ICU stay to vary with the number of days in the ICU when infected. To do this we changed the regression equation (1) to

$$r_{ij} = \mathbf{x}_i^T \mathbf{b}_j + x_i' c_{jd} I(d_i = d), \qquad i = 1, \ldots, n,\ j = 2, 3, \tag{2}$$

where x_i' is a time-dependent exposure and d_i is the days since ICU entry for row i of the data. The parameter c_{jd} is then the effect of x_i' on day d, and is estimated separately for discharge ($j = 2$) and death ($j = 3$). We estimated the c_{jd} parameters as Normally distributed random effects given by

$$c_{jd} \sim \mathrm{N}(\mu_c, \sigma_c^2).$$

Because the number of patients not discharged or dead becomes small as d becomes large, we truncated the time-varying intercept after 21 days. So for $d \geq 22$, $c_{jd} = c_{j22}$.

Equation (2) can also be used to model a lagged effect for a time-dependent covariate if d is defined as the number of days since x_i' equalled X. This lagged effect allows the effect of the covariate to change after a specific event ($x_i' = X$). In this analysis, we allowed the effect of MRSA infection to vary with the time since first infected, because we were interested in whether the effect of the MRSA infection on discharge waned with increasing time since infection.

Estimating the extra length of stay

To calculate the excess length of stay due to infection on day d we subtracted the survivor functions from the day of infection onwards using

$$E(\text{excess LOS}|\text{MRSA on day } d) = \sum_{t=d}^{m} S(t|\text{MRSA on day } d) - S(t|\text{no MRSA}).$$

We only evaluate the sum up to some limit m, as for large values of t the survivor functions become very small. In this analysis we use a limit of $m = 21$ days. We estimated the survivor function at day t by multiplying the probabilities of staying from day 1 up to day t,

$$S(t) = \pi_{j1}(t' = 1, \mathbf{x}_j) \times \pi_{j1}(t' = 2, \mathbf{x}_j) \times \ldots \times \pi_{j1}(t' = t, \mathbf{x}_j)$$

where $\mathbf{x}_j$ is a set of covariates (we have used a different notation to that above to emphasise the dependence of the probability on time and the covariates). We estimated the excess length of stay for an infection occurring during each of the first 21 days in ICU, and for three different Day 1 APACHE II scores: 10, 20 and 30, reflecting a range of morbidity.

Model fitting and building

We fitted a number of different models using the model extensions detailed above using the same set of exposures for each model (Table 1). We selected the best fitting model using the deviance information criterion (DIC). A dif-

ference in DIC of 10 is considered substantial (http://www.mrc-bsu.cam.ac.uk/bugs/winbugs/dicpage.shtml).

The models were fitted in a Bayesian framework using the WinBUGS package (16), using vague priors for all unknown parameters. We used a vague Normal prior with zero mean and variance of 1000 for all regression parameters, and a gamma prior with a shape and inverse scale parameter of 0.001 for all inverse-variance parameters.

Missing data

As shown in Table 1, there was some missing data for the TISS score, and just three missing scores for the day 1 APACHE II. The TISS scores were most often missing on the day of a patient's discharge or death. To prevent these important days from being lost from the analysis we imputed the TISS scores using a random effect for each admission given by

$$\begin{aligned} x_{ik} &\sim \mathrm{N}(\mu_i, \sigma_w^2), \qquad i = 1, \ldots, n,\ k = 1, \ldots, m_i, \\ \mu_i &\sim \mathrm{N}(\mu, \sigma_b^2), \qquad i = 1, \ldots, n, \end{aligned}$$

where m_i is the number of days observed for subject i, μ_i is the mean score for each subject, and σ_b^2 and σ_w^2 estimate the between- and within-subject variance, respectively. This imputation was made in WinBUGS in tandem with the estimation of the parameters governing discharge and death. In order to keep the regression model and imputation separate we used the

"cut" function (17).

RESULTS

There were 4569 separate admissions leading to 44,505 days spent on the ICU. Lengths of patient stays ranged from 0 to 363 days, and were highly skewed with a mean (SD) of 8.8 (14.1) days and a median (IQR) of 4 (2, 11). MRSA was cultured from 864 patients of which 335 developed an MRSA infection which led to 6696 infected days: 15.0% of the total patient-days on the ICU. One-hundred and six (31.6%) of the 335 admissions who developed an MRSA infection died, compared with 916 out of 4234 (21.6%) of the admissions without an MRSA infection.

We fitted seven different models to the data and they are compared using the DIC in Table 2. Model I had no random intercepts or time-dependent covariates (see table 1 for list of exposures). Model II built on Model I by including a time-dependent intercept to account for the fact that risk of death or discharge might vary during a patient's stay on ICU. This decreased the DIC by 710 indicating a substantially improved model fit. The probabilities that a patient stayed, died or was discharged during the first 21 days in ICU are shown in Figure 1. The most notable feature is the change from a low probability of discharge of 0.08 on the day of admission (day 0) to a peak in discharge probability on day 1 of 0.29. After day 1 the daily probability of discharge gradually decreases with increasing stay. The probability of death

is small on any day, ranging between 0.01 and 0.02.

Model III extended Model II by adding random admission-specific intercepts to account for between-admission variation not explained by other covariates. This addition improved the model fit greatly as the DIC decreased by 520 (Table 2).

Model IV built on Model III by allowing the effect of MRSA infection on death and discharge to vary depending upon when it occurred over the first 21 days after ICU admission. This addition worsened the model fit as the DIC increased by 10 (Table 2). The time-dependent MRSA estimates for the daily relative risk ratios of discharge and death are shown in Figure 2. MRSA infection was associated with a decreased risk of discharge, and this decrease was similar regardless of whether the MRSA infection occurred early or late in a patient's stay. The mean relative risk ratio of discharge after becoming infected with MRSA was 0.78 (95% CI: 0.62, 0.97) relative to an MRSA free patient. Developing an MRSA infection had little direct effect on the risk of death (RRR = 1.12, 95% CI: 0.88, 1.38).

Model V extended Model III by allowing the effect of MRSA infection on death and discharge, to vary from the day the infection started to examine whether the effect waned over time. This addition improved the model fit as the DIC decreased by 30 (Table 2). The time-dependent MRSA estimates for discharge and death are shown in Figure 3. The mean relative risk ratio of discharge after an MRSA infection was 0.76 (95% CI: 0.60, 0.97) relative to an MRSA free patient. There was a slight decrease in the effect of an

MRSA infection with increasing time since infection: during the first five days after infection the mean relative risk ratio of discharge was 0.73, while 10 days after infection it was 0.77. Again, there was little evidence that MRSA infection had a direct effect on the risk of death (RRR = 1.16, 95% CI: 0.95, 1.37). This result is considered further in the discussion.

Models I–V all adjusted for the daily TISS score. We were concerned that TISS score could be affected not only by the underlying severity of patient illness but also by MRSA infection. If this were the case, adjusting for the daily TISS score would bias the estimate of the effect of MRSA on death and discharge (18), possibly causing the model to miss a true association between MRSA and death. We therefore fitted two more models (Model VI and VII). Model VI was as Model V, but adjusted for day 1 TISS only, instead of daily TISS. The effect of MRSA infection on death changed little, but the relative risk ratio of discharge strengthened to 0.69 (95% CI: 0.55, 0.84). Model VII was as model III, but without any adjustment for TISS score, the effect on discharge was similar, RRR = 0.68 (95% CI: 0.54, 0.82).

The risk ratios of death and discharge relative to staying from Model VII are shown in Table 3. The strongest reduction in the risk of discharge was associated with ventilation, while haemofiltration was associated with the strongest increase in the risk of death. There was little evidence of difference in the risks of discharge or death between the five specialty categories. The risk of discharge varied by day of the week, being significantly lower on the weekend compared to Wednesday.

Figure 4 shows the relative risk ratio of discharge and death after an MRSA infection for the seven different models. For models IV to VI the effect plotted is the mean over all times. Including a time-dependent intercept greatly changed the effect of MRSA infection as shown by the differences in relative risk ratios between models I and II. Adding a random intercept meant that the effect of MRSA significantly decreased the risk of discharge (Model III compared with Model II). Models IV and V had similar mean relative risk ratios to Model III, but with slightly wider credible intervals. Models VI and VII showed the strongest reduction in the risk of discharge risk after MRSA infection.

Figure 5 shows the mean excess length of stay due to an MRSA infection according to day 1 APACHE scores. For an infection on day 1, a patient with an APACHE II score of 10 would have a mean extra length of stay of 0.3 days (95% CI: 0.1, 0.5). A sicker patient with an APACHE II score of 30 would a longer stay of 1.2 days (95% CI: 0.5, 2.0).

DISCUSSION

MRSA infection had little direct effect on risk of death (Figure 4; model VII RRR = 1.14, 95% CI: 0.93, 1.39). However, all the models considered indicated that infection indirectly contributed to increased mortality. This occurs because patients with MRSA infections tend to stay longer in the ICU, and each day in the ICU has an associated mortality risk. Such an indirect

link has also been found between nosocomial pneumonia infection and death (19).

All seven multinomial models considered found that MRSA infection decreased the risk of discharge. While the magnitude of this effect varied considerably between models, the three models which gave (by some margin) the best fits to the data (III, IV and V) all yielded remarkably similar estimates: patients with MRSA infections had a relative risk of discharge (compared to staying) that was about 20% lower than that for patients without MRSA infections.

Additional stay attributed to the MRSA infection was found to be higher for sicker patients (as measured by the APACHE II score) and for infections occurring earlier in the ICU stay (Figure 5). The mean excesses are for all admissions, so infections occurring at later times have a much smaller attributable cohort and so cause less overall excess stays. Sicker patients have less physical reserve and may be less able to cope with an infection, hence their increased length of stay after infection compared with healthier patients.

One potential drawback of our approach is the retrospective nature of the infection diagnosis based on SIRS criteria, antibiotic start and culture of MRSA from a clinical site. This led to a diagnosis of MRSA infection in 335 (39%) of 864 patients colonized with MRSA, which although high is likely to be explained by the particular virulence of MRSA compared with other hospital bacteria and the hyperinvasive nature of the TW MRSA strain that

was circulating on the ICU at that time (9).

Our approach to estimating additional length of stay caused by hospital infections overcomes the pitfalls that affect much of the literature, and we believe it should therefore provide more reliable estimates. However, an alternative analytical approach would have been to use a multistate model (14, 19, 20).

Such multistate models can be used to model the flow of patients through a set of defined states. For example, patients may start in the "hospital entry" state, some may subsequently move to an "infected" state, and then to a "discharged" state. However, an important advantage of the approach we used is the relative ease of incorporating lagged covariates and random effects. Nonetheless, the two approaches have similarities, and the multinomial model used in this paper can be thought of as a discrete time analogue of a multistate model (when comparable models could be fitted, they were in fact found to give very similar results). One potential drawback of the multinomial model is that the data must be equally spaced, and in our case all times were rounded to the nearest day. However, the model could be readily extended to smaller time-steps if required and data were available.

The aim of this work was to estimate the degree to which MRSA infection causes increased length of stay, rather than to simply document associations. This raises important issues about identifying confounders. Controlling for the time-dependent exposure of daily TISS score gave a better fit to the data (Table 2), but it is also likely that MRSA infection will have an effect on the

TISS score, particularly through prompting new antibiotic starts, catheter insertions and treatment with vasoactive agents. Therefore adjusting for the daily TISS score in the model would not only be unnecessary, it would be potentially actively harmful and introduce bias (18). This consideration motivated Models VI (which includes only the day 1 rather than the daily TISS score) and Model VII (without any TISS score), which are therefore not vulnerable to this problem. These models showed an increased effect of MRSA infection on discharge compared with the other models with time-dependent covariates (Figure 4) (18).

Being able to compare the fit of the models using the DIC is one of the advantages of using a Bayesian framework (Table 2). One surprising result was the reduction in the effective number of parameters from 2208 for Model V to 1604 for Model VI, when the only change was fitting TISS as a time-independent, instead of time-dependent, covariate (hence we might have expected a reduction of around 40 parameters). The large reduction in the effective number of parameters is due to the day 1 TISS score explaining a lot of the between-admission variation in the risk of discharge and death. This variation is no longer modelled by the random subject intercept, and so many less parameters are needed. Nonetheless, despite the poorer fit to data as measured by the DIC, knowledge of the likely causal pathways suggests that models VI and VII would give the most reliable estimate of the impact of MRSA infection on length of stay.

References

[1] Vincent JL, Bihari DJ, Suter PM, et al. The prevalence of nosocomial infection in intensive care units in Europe. results of the European prevalence of infection in intensive care (EPIC) study. EPIC international advisory committee. *JAMA* 1995;274(8):639–44.

[2] Harbarth S, Sax H, Gastmeier P. The preventable proportion of nosocomial infections: an overview of published reports. *J Hosp Infect* 2003; 54:258–66.

[3] Graves N. Economics and preventing hospital-acquired infection. *Emerg Infect Dis* 2004;10(4):561–566.

[4] Graves N, McGowan J. Nosocomial infection, the deficit reduction act and incentives for hospitals. *JAMA* 2009;in press.

[5] Faddy M, Graves N, Pettitt A. Modeling length of stay in hospital and other right skewed data: Comparison of phase-type, gamma and log-normal distributions. *Value in Health* 2008;9999(9999). 10.1111/j.1524-4733.2008.00421.x.

[6] Graves N, Weinhold D, Roberts J. Correcting for bias when estimating the cost of hospital acquired infection: An analysis of lower respiratory tract infections in non-surgical patients. *Health Econ* 2005;14(7):755–761.

[7] Diggle P, Heagerty P, Liang KY, et al. *Analysis of Longitudinal Data*. Oxford Statistical Science Series. USA: Oxford University Press, 2nd ed., 2002.

[8] Keene A, Cullen D. Therapeutic intervention scoring system: update 1983. *Crit Care Med* 1983;11:1–3.

[9] Edgeworth J, Yadegarfar G, Pathak S, et al. An outbreak in an intensive care unit of a strain of methicillin-resistant staphylococcus aureus sequence type 239 associated with an increased rate of vascular access device-related bacteremia. *Clin Infect Dis* 2007;44:493–501.

[10] Batra R, Eziefula A, Wyncoll D, et al. Throat and rectal swabs may have an important role in MRSA screening of critically ill patients. *Intensive Care Med* 2008;34:1703–6.

[11] American college of chest physicians/society of critical care medicine consensus conference: Definitions for sepsis and organ failure and guidelines for the use of innovative therapies in sepsis. *Crit Care Med* 1992; 20(6):864–874.

[12] Dobson AJ, Barnett AG. *An Introduction to Generalized Linear Models*. Texts in Statistical Science. Boca Raton, Flo.: Chapman & Hall/CRC, 3rd ed., 2008.

[13] Hedeker D, Gibbons RD. *Longitudinal Data Analysis*. Wiley Series in Probability and Statistics. New Jersey: Wiley, 2006.

[14] Putter H, Fiocco M, Geskus R. Tutorial in biostatistics: competing risks and multi-state models. *Statistics in Medicine* 2007;26(11):2389–2430.

[15] D'Agostino RB, Lee ML, Belanger AJ, et al. Relation of pooled logistic regression to time dependent cox regression analysis: The Framingham heart study. *Statistics in Medicine* 1990;9(12):1501–1515.

[16] Spiegelhalter D, Thomas A, Best N, et al. WinBUGS version 1.4.3 user manual, MRC Biostatistics Unit, Cambridge, 2007.

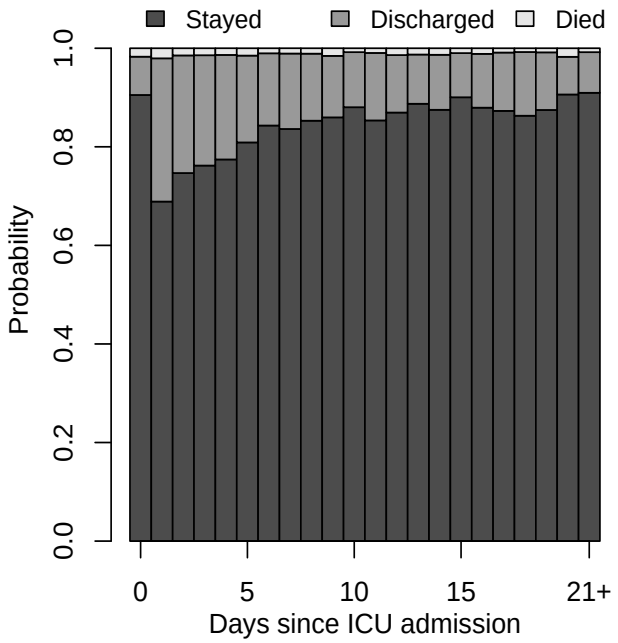
[17] Carrigan G, Barnett A, Dobson A, et al. Compensating for missing data from longitudinal studies using WinBUGS. *Journal of Statistical Software* 2007;19.

[18] Hernán MA, Hernández-Daz S, Werler MM, et al. Causal knowledge as a prerequisite for confounding evaluation: An application to birth defects epidemiology. *American Journal of Epidemiology* 2002;155(2):176–184.

[19] Wolkewitz M, Vonberg RP, Grundmann H, et al. Risk factors for the development of nosocomial pneumonia and mortality on intensive care units: Application of competing risks models. *Critical Care* 2008; 12(2):http://ccforum.com/content/12/2/R44.

[20] Wolkewitz M, Beyersmann J, Gastmeier P, et al. Modeling the effect of time-dependent exposure on the occurrence of death in icu population. *Int Care Med* 2009;in press.

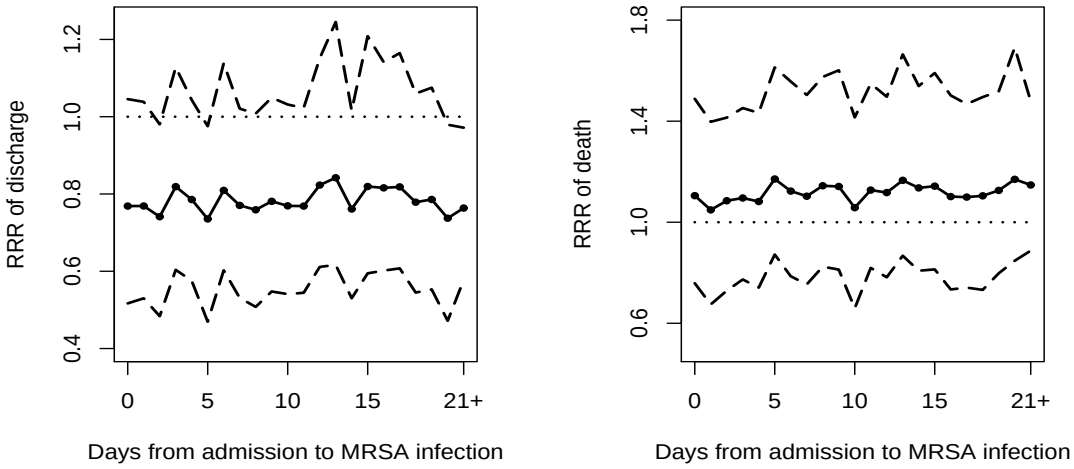
Figure 1: Probability that a patient stayed, was discharged or died by days since ICU admission (estimates from model II)



intensive care unit (ICU)

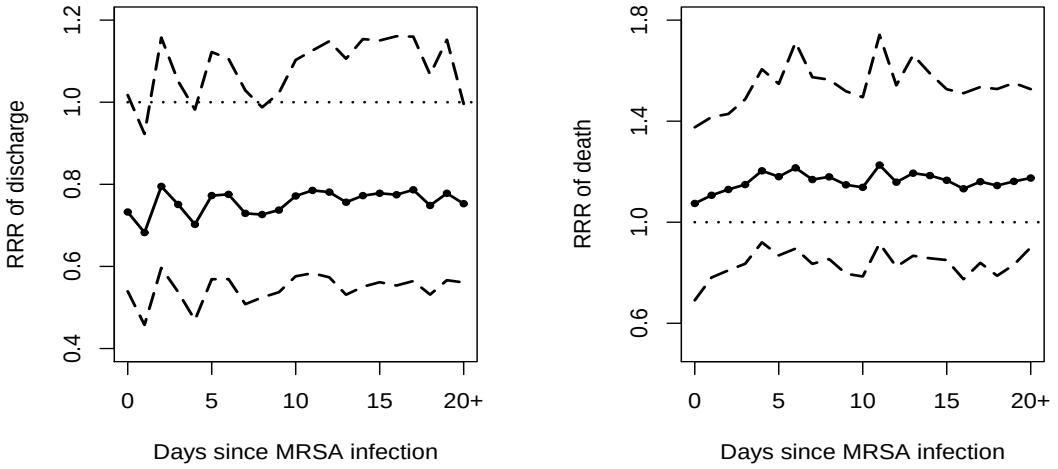
Mean probabilities from model II for a female patient aged 60, with a day 1 APACHE II of 20, daily TISS of 40, who is not ventilated or filtrated, in the specialty category of Medicine + Acute Renal Failure + Cardiology

Figure 2: Daily relative risk ratio of discharge and death from ICU relative to staying after an MRSA infection dependent on the time delay between admission and infection (estimates from model IV)



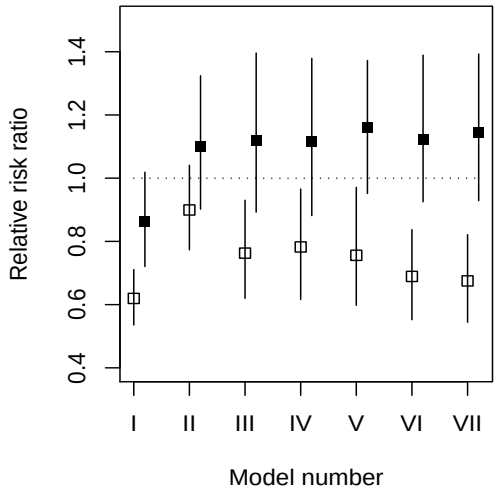
intensive care unit (ICU), Methicillin-resistant *Staphylococcus aureus* (MRSA), relative risk ratio (RRR) Mean relative risk ratio and 95% credible interval. Horizontal line at relative risk ratio of 1.

Figure 3: Daily relative risk ratio of discharge and death from ICU relative to staying after an MRSA infection by days since infection (estimates from model V)



intensive care unit (ICU), Methicillin-resistant *Staphylococcus aureus* (MRSA), relative risk ratio (RRR) Mean relative risk ratio and 95% credible interval. Horizontal line at relative risk ratio of 1.

Figure 4: Overall daily relative risk ratio of discharge and death from ICU relative to staying after developing an MRSA infection by model number



intensive care unit (ICU), Methicillin-resistant *Staphylococcus aureus* (MRSA)

Mean relative risk ratio and 95% credible interval. Open square for mean relative risk ratio of discharge, closed square for mean relative risk ratio of death. Horizontal line at relative risk ratio of 1.

Figure 5: Mean excess length of stay due to an MRSA infection by day of infection and Day 1 APACHE II score (estimates from model VII)

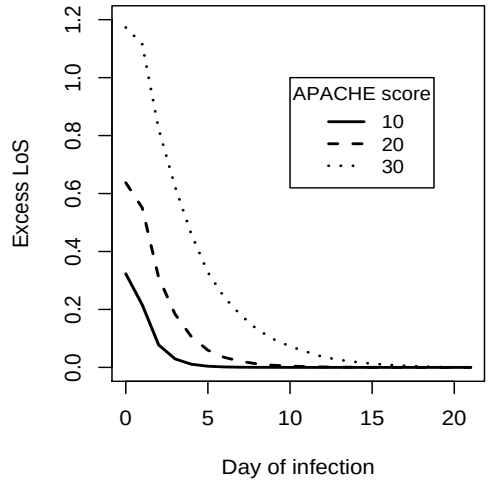


Table 1: Exposures Used and Descriptive Statistics

Covariate (category)	Time-dependent	Type	Missing	Statistics[†]
Age	No	Continuous	0	60 (17.5)
Sex (male)	No	Binary	0	2815 (61.6)
Speciality	No	Nominal	0	
Day of week	Yes	Nominal	0	
Day 1 APACHE II	No	Continuous	3 (0.1%)	17.2 (8.0)
TISS	Yes	Continuous	3844 (8.6%)	40.7 (13.4)
Ventilation (yes)	Yes	Binary	0	33,015 (74.2)
Haemofiltrated (yes)	Yes	Binary	0	7408 (16.6)
MRSA infection[‡] (yes)	Yes	Binary	0	6696 (15.0)

acute physiology and chronic health evaluation (APACHE), credible interval (CI), Methicillin-resistant *Staphylococcus aureus* (MRSA), therapeutic intervention scoring system (TISS)

[†] Mean (standard deviation) for continuous covariates, number (% of patients) for sex and number (% of days) for binary time-dependent covariates

[‡] MRSA carried forward from day of infection to all subsequent days

Table 2: Comparing Different Models Using the Deviance Information Criterion

Model	Effective Number of parameters	Mean deviance	DIC	Difference in DIC[†]
I) No time-dependent covariates or random intercepts	333.0	27,070	27,400	
II) Model I + time-dependent intercept	382.6	26,310	26,690	710
III) Model II + random admission intercepts	2148	24,020	26,170	520
IV) Model III + time-dependent MRSA	2117	24,060	26,180	−10
V) Model III + lagged time-dependent MRSA	2208	23,930	26,140	30
VI) Model V − daily TISS + day 1 TISS	1640	25,370	27,370	−1230
VII) Model III − daily TISS	1649	25,750	27,400	−3340

deviance information criterion (DIC)

[†] Lower numbered model minus higher. A lower DIC indicates a better model fit; a difference in DIC of 10 is considered substantial.

Table 3: Relative Risk Ratios (and 95% credible intervals) of Daily Discharge and Death Relative to Staying. Estimates from Model VII.

	Discharge		Death	
Variable (units)	RRR	95% CI	RRR	95% CI
Age (10 year increase)	0.93	0.90, 0.96	1.13	1.08, 1.18
Sex (Male vs Female)	1.12	1.01, 1.25	0.81	0.71, 0.93
Day 1 APACHE II (5 point increase)	0.66	0.63, 0.69	1.32	1.25, 1.40
Ventilated (yes vs no)	0.05	0.04, 0.06	1.63	1.36, 1.99
Heamofiltration (yes vs no)	0.30	0.25, 0.36	1.80	1.56, 2.08
MRSA infection (yes vs no)	0.68	0.54, 0.82	1.14	0.93, 1.39
Speciality† (Surgery)	0.87	0.74, 1.01	1.07	0.89, 1.29
Speciality† (Cardiothoracic surgery)	0.94	0.80, 1.09	1.02	0.85, 1.23
Speciality† (Orthapaedics)	0.84	0.55, 1.24	1.21	0.67, 1.95
Speciality† (ITU referrals)	0.93	0.78, 1.11	1.14	0.93, 1.39
Day of the week‡ (Monday)	0.89	0.77, 1.05	0.85	0.68, 1.07
Day of the week‡ (Tuesday)	0.94	0.81, 1.11	0.95	0.76, 1.18
Day of the week‡ (Thursday)	1.00	0.85, 1.17	0.91	0.72, 1.14
Day of the week‡ (Friday)	1.06	0.92, 1.24	0.93	0.73, 1.17
Day of the week‡ (Saturday)	0.72	0.62, 0.86	0.90	0.72, 1.13
Day of the week‡ (Sunday)	0.63	0.53, 0.75	0.98	0.78, 1.24

acute physiology and chronic health evaluation (APACHE), credible interval (CI), Methicillin-resistant *Staphylococcus aureus* (MRSA), relative risk ratio (RRR), therapeutic intervention scoring system (TISS)

† Reference category = Medicine + Acute Renal Failure + Cardiology

‡ Reference category = Wednesday